Using GPU Nodes

GPU Node Reservation: Until February 28, 2022, thirty Cascade Lake-V100 GPU nodes (cas_gpu) are reserved for a special project. During this time, we recommend using the Skylake-V100 GPU nodes (sky_gpu) for your jobs requiring V100 GPUs.
A graphics processing unit (GPU) is a hardware device that can accelerate some computer codes and algorithms. This article describes the two types of GPU nodes currently available at NAS. For information about running your PBS jobs on the GPU nodes, see Requesting GPU Resources. For information about developing code for the GPU nodes, see GPU Programming.

For more general information on GPUs, see the GPGPU article on Wikipedia and the NVIDIA website, which has information for developers.

If you have an application that can take advantage of GPU technology, you can use one of the NAS GPU models:

- model=san_gpu

  64 nodes total. Each node contains two Sandy Bridge CPU sockets and one NVIDIA Kepler K40 GPU coprocessor (card) connected via PCI Express bus.

  The Sandy Bridge sockets used in the **san_gpu** nodes and those in the other Pleiades Sandy Bridge nodes are both Intel Xeon E5-2670. However, the size of memory in the sockets of the **san_gpu** nodes is double that of the other Pleiades Sandy Bridge nodes.
- model=sky_gpu

  19 nodes total. There are two different configurations:

  - 17* nodes, each containing two Skylake CPU sockets and **four** NVIDIA Volta V100 GPU cards, where the CPUs and GPUs are connected via PCI Express bus
  - Two nodes, each containing two Skylake CPU sockets and **eight** NVIDIA Volta V100 GPU cards, where the CPUs and GPUs are connected via PCI Express bus

  The Skylake sockets used in the **sky_gpu** nodes are Intel Xeon Gold 6154 while those in the Electra Skylake nodes are Intel Xeon Gold 6148. The main differences in these two configurations (**sky_gpu** vs. Electra Skylake) are: (1) number of cores - 18 vs. 20; (2) CPU clock speed - 3.0 vs. 2.4 GHz; (3) L3 cache size - 24.75 vs. 27.5 MB; (4) memory size - 384 vs. 192 GB; and (5) number of UPI links - 3 vs. 2.

  * Three of the 17 nodes are reserved for a special project; 14 nodes are available for use.
- model=cas_gpu

  38* nodes total. Each node contains two Cascade Lake CPU sockets and four NVIDIA Volta V100 GPU cards, where the CPUs and GPUs are connected via PCI Express bus.

  The Cascade Lake sockets used in the cas_gpu nodes are Intel Xeon Platinum 8268 while those in the Aitken Cascade Lake nodes are Intel Xeon Gold 6248. The main differences in these two configurations (cas_gpu vs Aitken Cascade Lake) are: (1) number of cores - 24 vs 20; (2) CPU clock speed - 2.9 vs. 2.5 GHz; (3) L3 cache size - 35.75 vs. 27.5 MB; and (4) memory size (384 vs 192 GB).

  * Four of the 38 nodes are reserved for a special project. 34 nodes are available for public use.

The SBU rates for the GPU nodes are:

- **san_gpu**: 0.94
- **sky_gpu** with four V100s: 9.82
- **sky_gpu** with eight V100s: 15.55
- **cas_gpu** with four V100s: 9.82

Node Details			
	san_gpu	**sky_gpu**	**cas_gpu**
Architecture	In-house	Apollo 6500 Gen10	Apollo 6500 Gen10
Total # of Nodes	64	19	38
Host name	r313i[0-3]n[0-15]	r101i0n[0-11,14-15], r101i1n[0-2] and r101i0n[12-13]*	r101i2n[0-17], r101i3n[0-15] and r101i4n[0-3]
CPU Host			
Processors Model	8-core Xeon E5-2670	18-core Xeon Gold 6154	24-core Xeon Platinum 8268
# of CPU Cores/Node	16	36	48
CPU-Clock	2.6 GHz	3.0 GHz	2.9 GHz
Maximum Double Precision Floating Point Operations per cycle per core	8	32	32
Total CPU Double Precision Flops/Node	332.8 GFlops	3,456 GFlops	4,454 GFlops
Memory/Node	64 GB (DDR3)	384 GB (DDR4)	384 GB (DDR4)
GPU			
Device Name	Tesla K40m	Tesla V100-SXM2-32GB	Tesla V100-SXM2-32GB
Clock Rate (Base/Boost)	745 MHz/875 MHz	1290 MHz/1530 MHz	1290 MHz/1530 MHz
# of coprocessors/Node	1	4 (17 nodes) or 8 (2 nodes)	4
# of Single Precision CUDA Cores/coprocessor	2880	5120	5120
# of Double Precision CUDA Cores/coprocessor	960	2560	2560
# of Tensor Cores/coprocessor	N/A	640	640
Total GPU Double Precision Flops/coprocessor (Base/Boost)	1.43 TFlops/1.68 TFlops	6.6 TFlops/7.8 TFlops	6.6 TFlops/7.8 TFlops
Total GPU Double Precision Flops/Node (Base/Boost)	1.43 TFlops/1.68 TFlops	26.4 TFlops/31.2 TFlops for 4 V100 and 52.8 TFlops/62.4 TFlops for 8 V100	26.4 TFlops/31.2 TFlops
L2 Cache	1536 KB	6144 KB	6144 KB
Global Memory/coprocessor	12 GB (GDDR5)	32 GB (HBM2)	32 GB (HBM2)
Memory Clock Rate	3004 MHz	877 MHz	877 MHz
Memory Bus Width	384 bits	4096 bits	4096 bits
Memory Bandwidth	288 GB/s	900 GB/s	900 GB/s

GPU-to-GPU Communication	N/A	SXM2 mezzanine connector with NVLink 2.0 with 300 GB/s bi-directional bandwidth	SXM2 mezzanine connector with NVLink 2.0 with 300 GB/s bi-directional bandwidth
PGI Compiler Option	-ta=tesla:cc35	-ta=tesla:cc70	-ta=tesla:cc70
Inter-node Network			
IB Device on Node	1 FDR 2-port card	2 EDR 100 Gb 2-port cards	2 EDR 100 Gb 2-port cards
Local Disk			
SSD/node	N/A	** 3.2 TB raw, 1.6 TB usable (mirrored) (2 x 1.6 TB SAS SSD connected via an internal smart Raid card)	** 3.2 TB raw, 1.6 TB usable (mirrored) (2 x 1.6 TB SAS SSD connected via an internal smart Raid card)

* r101i0n[12-13] are for the two nodes containing eight V100 cards.

** Not yet available for public usage, pending configuration decisions.

For more hardware details, access a GPU node through a PBS session and run one of the following commands:

```
For CPU info:

cat /proc/cpuinfo

For host memory info:

cat /proc/meminfo


For GPU info:

/usr/bin/nvidia-smi -q

or load a PGI compiler module, for example, comp-pgi/20.4, and run the command pgaccelinfo.


module use -a /nasa/modulefiles/testing
module avail comp-pgi
---- /nasa/modulefiles/testing ----
comp-pgi/17.10 comp-pgi/18.10 comp-pgi/18.4 comp-pgi/19.10 comp-pgi/19.5 comp-pgi/20.4
---- /nasa/modulefiles/sles12 -----
comp-pgi/16.10 comp-pgi/17.1

module load comp-pgi/20.4
pgaccelinfo
```

Currently, no direct communication exists between a GPU on one node and a GPU on another node. If such communication is required, the data must go from the GPU to the CPU via the PCI Express bus, and from one CPU to another via InfiniBand network using MPI.

Article ID: 298
Last updated: 18 Nov, 2021
Revision: 75
Machine Learning -> Machine Learning at NAS -> Using GPU Nodes

https://www.nas.nasa.gov/hecc/support/kb/entry/298/